An Improved Framework Using LDA and LSTM with RSO for Intrusion Detection in IoT

Hussein Ali Manji Nasrawi[1], Hasanain Salim Kadhim[2], Rifaq Adil Majeed[2] and Hussain Ali Hussain[2]
[1]*Department of Computer Science, University of Kufa, 54001 Najaf, Iraq*
[2]*Department of Education in Najaf Governorate, Iraqi Ministry of Education, 54001 Najaf, Iraq*
husseina.alnesrawi@uokufa.edu.iq, hsiraq1985@gmail.com, refaqadel@gmail.com, ahusseina86@gmail.com

Keywords: Wireless Sensor Networks, Intrusion Detection, Long Short-Term Memory (LSTM), Rat Swarm Optimization (RSO), Feature Selection (LDA).

Abstract: This paper proposes a lightweight and efficient intrusion detection framework for Internet of Things (IoT) and Wireless Sensor Network (WSN) environments by integrating Linear Discriminant Analysis (LDA), Long Short-Term Memory (LSTM) networks, and Rat Swarm Optimization (RSO). The proposed approach first applies LDA to reduce high-dimensional network traffic data into a compact and discriminative feature space, improving class separability while reducing computational complexity. Subsequently, an LSTM-based deep learning model is employed to capture temporal dependencies in network traffic for multi-class intrusion classification. To enhance model performance and stability, RSO is utilized to optimize critical LSTM hyperparameters, including learning rate, batch size, and hidden units, enabling an effective balance between exploration and exploitation during training. The framework is evaluated on two benchmark datasets, CIC-IDS2018 and TON_IoT, representing diverse and realistic IoT attack scenarios. Experimental results demonstrate that the proposed model achieves high detection performance, with accuracy exceeding 99%, along with strong precision, recall, and F1-scores, and a low false alarm rate. Furthermore, ROC-AUC and PR-AUC values confirm robust discriminative capability under imbalanced data conditions. Deployment experiments on edge devices such as Raspberry Pi 4 and NVIDIA Jetson Nano show low inference latency and moderate resource consumption, validating the suitability of the proposed framework for real-time intrusion detection in resource-constrained IoT environments.

1 INTRODUCTION

Wireless Sensor Networks (WSNs) have played a significant role in the realization of cyber-physical systems in applications such as healthcare, smart agriculture, industrial automation and environmental monitoring [1], [2]. These are micro-sensors, a small device which is power node and equipped with sensing, processing and communication functionalities. However, their bandwidth capabilities and power resources are also quite constrained. privacy preservation method to recover information service) sensor networks However, a large proportion of these networks are forced to work in open or hostile environments which make them extremely vulnerable to different kinds of security attacks [3]. One method to detect (and prevent) this behaviour is via Intrusion Prevention Systems (IPS), which acts as our first layer of defense, in an effort to catch and halt the attack before it has done too much damage [4], [5]. However, the classical IPS suffer from the tradeoff between accuracy and energy efficiency, being in too strict-latency, low-energy and small memory constrained WSNs [6]. Every advance in AI brings fresh ways to address these problems. Deep learning, including Long Short-Term Memory (LSTM) network and Convolutional Neural Network (CNN), are effective models for large-scale, complex higher-dimensional network traffic data [7], [8]. To make the method computationally more effective, we can apply Feature Selection (FS) techniques like LDA for extracting compact and discriminating features that are also human interpretable [9]. On the other hand, swarm intelligence algorithms such as RSO offer a set of heuristics to fix hyperparameters at an acceptable level of exploration/exploitation in high-dimensional search spaces [10].

Inspired by these studies, we propose in this paper an integrated IPS framework of LDA and RSO with adding LSTM. The above framework LDA reduces the dimension of input features, RSO aims at optimizing parameters of LSTM, and finally LSTM performs an effective temporal anomaly classification. Its performance is evaluated on CIC-IDS 2018 and TON_IoT datasets, which represent

real-world and diverse attack behaviors. The contributions of this work are fourfold: (i) using LDA for achieving efficiency while maintaining interpretability, (ii) improving the performance of LSTM by optimizing the RSO method based on it, (iii) validating across heterogeneous datasets supporting its generalizability, and iv) examining real-time feasibility with experiments on low-power edge devices. The remainder of this paper is structured as follows: Section 2 reviews related IPS approaches, Section 3 presents the proposed methodology, Section 4 describes the experimental setup, Section 5 reports and discusses the results, and Section 6 concludes with future research directions.

2 RELATED WORK

Widespread and growing applications of wireless sensor networks (WSNs) across the critical infrastructure sectors, ranging from healthcare and smart cities to industrial automation, have made urgent the need to address newfound cybersecurity problems. Despite their scalability and low operational cost, WSNs are vulnerable to a variety of attacks, including denial-of-service (DoS), unauthorized access, and data manipulation. Their attacks surface primarily because the networks themselves possess limited processing power, constrained and hard-to-manage energy resources, and a decentralized architecture that lacks security redundancy [11], [12]. In the past, traditional intrusion detection and prevention systems (IPS) for wireless sensor networks (WSNs) relied almost entirely on two types of detection methods: signature-based detection and statistical anomaly detection. Although these approaches are lightweight in terms of computational power, they are not very useful for identifying new or rapidly evolving threats, and this is the main reason why many researchers have recently begun favoring more intelligent and Machine Learning (ML)/Deep Learning (DL) based methods in order to provide adaptive and self-learning intrusion detection features [13]. Especially publicize, among them includes the LSTM network that also has demonstrated a huge power. LSTM, as a recurrent neural network, is especially conducive to learning the temporal dependencies of sequentially structured data-this makes it really well suited for carrying out on network traffic, the sort of time-series analysis that might expose long-latency intrusion signatures that a static model would miss-but as we've already gone over any flavor of deep learning models being trained on high-dimensional datasets could add potentially massive amounts of computational complexity and training time to your overheads that makes such an approach all but unfeasible outside of being carried out within a supercomputer [14]. This can be alleviated with dimensionality reduction methods such as LDA. The goal of LDA is to transform input matrix onto a space where classes are as separated as possible, at the same time the discriminant power and amount of redundancy along with corruption has been removed from transformed data. This has at least three advantages: First, you can then compute faster, second, you get a better interpretability and accuracy of your model and third it may be already some kind of useful for e.g. an WSN [15]. At the same time, this optimization approach has been exploited due to its capability to facilitate and improve the accuracy of detected models in deep learning models. Some bio-inspired metaheuristics were suggested in the context of optimization algorithms such as PSO, ACO and recently RSO [16]. These metaheuristics also serve as efficient tuners and locate the best hyperparameters for deep learning models with a trade-off balance between global exploration and local exploitation. Inspired by the adaptive foraging behavior of rats, RSO has been demonstrated to exhibit efficient convergence in high-dimensional complex search spaces-a very desirable property when addressing large and complex optimization problems such as training and testing large, deep learning models like the LSTM network [17]. Prior studies have demonstrated that LSTM-based IPS models perform well on the benchmarking datasets. One such reference dataset, that contains realistic network traffic as well as various attack flavors, is CIC-IDS 2018 [18]. This dataset includes not only the usual penetration-style attacks but also botnet and brute-force attempts; these datasets are more representative of the network traffic encountered in real-world environments. In contrast, many available datasets related to Industrial Control Systems (ICS) are outdated and were generated in controlled laboratory settings, often featuring a limited range of attack types [19]. New datasets, e.g., TON_IoT, offer much more diverse testing scenarios, hence the possibility of a more thorough assessment of current security-related approaches. Despite the advances, the literature is still lacking a unified approach that can efficiently be utilized for feature selection, deep learning, and dynamic optimization, particularly for the WSN. The majority of existing studies tend to overlook the critical role of hyperparameter tuning or rely on conventional approaches that are often computationally inefficient. To address these

limitations, in this paper, we propose a novel framework for real-time applications by considering three complementary parts. The first step, the LDA is used for feature reduction and selects only discriminative features to reduce computation cost. Second, a Rat Swarm Optimization (RSO) method is employed for adaptive hyper-parameter tuning, and it can achieve fast and stable optimization in real-time with high efficiency. Finally, the Long Short-Term Memory (LSTM) networks are used for sequential data and correctly classify the intrusion. These ingredients collectively contribute to making a strong and effective IPS solution that can balance performance vs practicality for resource-restricted WSN/IoT environments.

3 METHODOLOGY

In this paper, we have proposed a new IPS in which three main technologies are combined. The application of LDA ensures that the system focuses on important features. Threat classification over time is carried out by Long Short-Term Memory (LSTM) networks, and model hyperparameter learning with RSO is performed along the way. Working together in this manner, these components have proven to be not only effective for increasing the accuracy of detecting CML nodes; they also reduce the computing overheads, enabling this system to be lightweight enough so that it could run on real WSNS.

3.1 Overview of the Proposed Framework

The presented framework combines Linear Discriminant Analysis (LDA), Long Short Term Memory (LSTM) networks and Rat Swarm Optimization (RSO) to propose an intelligent intrusion detection prevention system that is lightweight. In general, the data processing flow begins from collecting and preprocessing a traffic dataset, which includes normalization and encoding as well as its definition in such a way to have non-disturbance properties. LDA maps the high-dimensional feature space into a low-dimensional subspace, preserving class discrimination and to minimize computational load. The features are down sampled and passed through the LSTM neural network to learn the temporal relationships and behavioral patterns in the network traffic. The parameter optimization scheme proposed in this work is employed to optimize the hyperparameters of the detection model using RSO which achieves a trade-off between adaptive exploration and exploitation behavior, resulting in better accuracy and faster convergence. The tool is designed to work as an IDS and IPS. In this mechanism, detection part detects anomalous network activity and prevention pattern takes measures of mitigating or blocking malicious traffic in a timely manner.

The two-fold functionality results in early threat detection and real-time response, which is important to preserve the security of IoT and WSN systems with limited resources.

With the integration of LDA-based dimensionality reduction, RSO-driven optimization and LSTM-based sequential modeling, the system striking a favorable trade-off between computation efficiency and detection accuracy. This approach in its entirety provides scalability, adaptability and readiness to be deployed in edge devices in practice [20].

3.2 Feature Selection Using (LDA)

The training cost and its overfitting might be severe problem to the high-dimensional input feature of CS datasets. We addressed this problem using LDA as a discriminant feature extractor. We have done this by determining a projection that maximizes the between-class scatter against within-class scatter ratio, so that class separability is increased and dimensionality of the data is reduced.

$$argmax_w \left(\frac{|W^T S_b W|}{|W^T S_w W|}\right), \quad (1)$$

where S_b and S_w represent the between-class and within-class scatter matrices, respectively, and W is the transformation matrix.

This transformation results in a representation where samples from different classes are far apart in the reduced feature space, which allows better classification.

Algorithm-1 shows the use of LDA in reducing the dimension of high-dimensional network traffic data, retaining important features used for class separation. In the case of Intrusion Prevention System (IPS) in WSNs, LDA as one of the important classifiers, attempts to achieve two ultimate goals: enhancing computing speed and improving classification accuracy. LDA does this by zeroing in on the two important aspects of the data. The first stage calculates average vectors of all classes within the data set and a global class average vector.

The averages offer an inexpensive description of the data, with the relevant qualitative characteristics

needed for the type of analysis they propose. They are developed to extract the class-specific and dataset-specific characteristic features in a more robust and consistent manner [21].

Algorithm 1: Linear Discriminant Analysis (LDA) for Feature Selection.

```
Input: X (n×d), y (n), k
for c in unique(y):
    μ_c ← mean(X[y==c], axis=0)
μ ← mean(X, axis=0)

S_w ← 0(d×d)
S_b ← 0(d×d)
for c in unique(y):
    Xc ← X[y==c]
    S_w ← S_w + Σ_{x∈Xc} (x-μ_c)(x-μ_c)^T
    n_c ← |Xc|
    S_b ← S_b + n_c(μ_c-μ)(μ_c-μ)^T

Solve S_b w = λ S_w w
Sort eigenpairs by λ descending
W_k ← first k eigenvectors (columns)
X' ← X W_k
return X'
```

Then, two essential scatter matrices are calculated to achieve the class separability in LDA. Inside class scatter matrix (S_w). It computes the compactness of all samples for every class by comparing the intra-class spread and inter-class variation. On the other hand, it can measure how well a pattern belongs to different classes that patterns of distinct classes are separated from each other, by calculating a distance based on divergence due to class mean deviation from its global average.

These matrices aim at constructing the optimal projection for class discrimination. LDA tries to find a projection matrix that maximizes the separation between classes and minimizes the in-class variance. From a mathematical point of view this amounts to solving a generalized eigenvalue problem, where the eigenvectors corresponding to the largest eigenvalues are those that maximize the separation between classes. Once the eigenvectors are computed, they are sorted according to their corresponding eigenvalues. The top k eigenvectors-those capturing the highest discriminatory power-are selected to form the final transformation matrix.

The original high-dimensional feature space is then transformed into a lower-dimensional one by applying the mutation matrix on the dataset. The reduced feature set X' is very informative, as it comprises most of the discriminative data needed for better intrusion detection with the reduction in overall computation overhead.

The condensed feature matrix is then supplied to the LSTM model for sequential pattern representation and intrusion identification. Through incorporating LDA as a part of the proposed framework, when compared to a purely supervised approach, the system enjoys faster training, lower risk of overfitting and simpler or more interpretable model generation and as such is particularly appealing for its potential use in resource limited WSN environment.

3.3 LSTM Network for Intrusion Classification

LSTM networks are selected for their ability to model temporal dependencies in sequential data, such as network traffic flows. The architecture includes:

- Input Layer. Receives LDA-reduced features.
- Two LSTM Layers. Each with 128 units.
- Dropout Layers. 30% rate to prevent overfitting.
- Output Layer. Dense layer with softmax activation for multi-class classification.

The model is trained with categorical cross-entropy loss and Adam optimizer with a fixed initial hyperparameters. This hyperparameters tuning is done using Rat Swarm Optimization (RSO). This procedure demonstrates the system-level design of LSTM based network IPS designed for WSNs, Incorporating model development through training and optimization for real-time operation.

Architecture We use two layers of LSTM with 128 hidden size to model long-term temporal dependency in traffic data, and dropout (0.3) are added to prevent overfitting [22]. The input features are preprocessed by LDA to make them relevant as well as to reduce dimensionality for better efficiency. Finally, a dense softmax output layer generates probability distributions over multiple classes (e.g., benign and various intrusion types), enabling robust multi-class intrusion detection.

The model is compiled with a loss function, categorical cross-entropy (useful for multi-category classification problems) and an Adam optimizer using default parameters for stable training. The trained extension model is then fine-tuned on the external dataset (CIC-IDS 2018 and TON_IoT) and we will monitor the performance of it using a validation set. But to achieve the best performance, important hyperparameters including batch size, learning rate or the number of hidden units are not manually assigned. Instead, we handle such a problem through an adaptive optimization algorithm

that uses RSO. The RSO rapidly explores the hyperparameter space by simulating the intelligent foraging of rats' colonies, and it coordinates exploration (searching new places) and exploitation (refining the places already known to deliver good results).

After finding the optimal hyperparameters using RSO such that maximum validation accuracy is achieved without overfitting, we retrain the LSTM model by employing the optimized configuration. This resulting model is then employed for deployment in real-time intrusion prevention, which can predict and prevent attacks in WSNs.

Algorithm 2: LSTM Model Training for Intrusion Prevention.

```
Input: X, Y, search spaces B={32,64,128}, H={64,128,256}, η∈[1e-5,1e-2]
Output: Trained LSTM model M*

function BuildModel(h, η):
    model := Input → LSTM(h) → Dropout(0.3) → LSTM(h) → Dropout(0.3) → Dense(C, softmax)
    compile(model, loss="categorical_crossentropy", optimizer=Adam(η), metrics=["accuracy"])
    return model

# Baseline
M0 := BuildModel(128, 1e-3)
score0 := TrainEval(M0, X, Y, batch=64, epochs=E0, early_stop=True)

# Fitness for RSO (lower is better)
function Fitness(b, η, h):
    M := BuildModel(h, η)
    val_score := TrainEval(M, X, Y, batch=b, epochs=E1, early_stop=True, use_val_split=True)
    return 1 - val_score

# Initialize swarm
Θ := {θ_i=(b_i, η_i, h_i)}_{i=1..P} # random samples
for i in 1..P: F_i := Fitness(θ_i)
θ* := argmin_θ F(θ)

# RSO main loop
for t in 1..T_max:
    for i in 1..P:
        θ_i := RSO_Update(θ_i, θ*, t) # continuous update
        η_i := clamp(η_i, 1e-5, 1e-2) # keep in range
        b_i := nearest(B, b_i) # map to discrete
        h_i := nearest(H, h_i)
        F_i := Fitness(θ_i)
        if F_i < Fitness(θ*): θ* := θ_i

# Final training with best hyperparams
(b*, η*, h*) := θ*
M* := BuildModel(h*, η*)
Train(M*, X, Y, batch=b*, epochs=E_final, early_stop=True)
return M*
```

3.4 Rat Swarm Optimization (RSO) for Hyperparameter Tuning

RSO is a swarm intelligence algorithm inspired by the social foraging behavior of rats. In this context, each rat represents a candidate hyperparameter configuration. The algorithm seeks to maximize model performance on a validation set.

Optimization Parameters:

- Batch size;
- Learning rate;
- Number of hidden units;
- Dropout rate.

Algorithm 3: Rat Swarm Optimization (RSO) for Hyperparameter Tuning.

```
Input: N, bounds [ℓ,u], fitness F(θ), T_max
Output: Gbest

# Init
for i = 1..N:
    θ[i] ← Uniform([ℓ,u])
    θ[i] ← snap_discretes(θ[i])
    Fval[i] ← F(θ[i])
Gbest ← argmin_i Fval[i]
for t = 1..T_max:
    ρ ← 1 - t / T_max
    for i = 1..N:
        if rand() < ρ:                # Exploration
            j ← random_index ≠ i
            θ_new ← θ[i] + α*randn(D) + β*(θ[j] - θ[i])
        else:                         # Exploitation
            θ_new ← Gbest + γ*(θ[i] - Gbest) + δ*randn(D)

        θ_new ← clip(θ_new, [ℓ,u])
        θ_new ← snap_discretes(θ_new)

        F_new ← F(θ_new)
        if F_new < Fval[i]:
            θ[i] ← θ_new
```

```
            Fval[i] ← F_new
            if F_new < F(Gbest):
                Gbest ← θ_new

return Gbest
```

Algorithm 3 employs Rat Swarm Optimization (RSO) to optimize hyperparameters of the LSTM-based IPS. RSO, inspired by rat foraging behavior, balances exploration and exploitation for effective search. A population of candidate solutions is initialized with hyperparameters: hidden units (16/32), learning rate (0.01/0.1/0.5), and batch size (10/20). Candidates are trained and evaluated on a validation set, with low-performing ones adjusted. RSO iteratively alternates between exploration (searching new regions) and exploitation (refining around the best solutions). Boundary checks enforce parameter limits, and the global best ($Gbest$) is updated until convergence or iteration limits. The LSTM is finally retrained with $Gbest$, producing an efficient IPS with superior detection accuracy for WSNs.

3.5 Integrated Workflow Implementation

The last system contains a simplified process. After pre-processing the data (normalization, clean up, encoding), LDA projects onto a lower-dimensional space while maintaining class information. The closer-to-the-beginning LSTM is set to set a performance baseline, before applying RSO for hyperparameter optimization and finding the best training settings. The trained model is retrained with the best parameters and utilized for online intrusion detection in WSNs. This pipeline is depicted in Figure 1 with highlights for its efficiency and suitability for edge deployment.

4 EXPERIMENTAL SETUP

We detail the experimental settings that were used to evaluate the proposed IPS in Section. The aim of the evaluation was (i) to assess the classification performance of the proposed LDA-LSTM-RSO pipeline in realistic WSN/IoT contexts and (ii) to evaluate its applicability at a practical level, with regard to both efficiency concerns and deployment scenarios. In order to remain true to reality, resource constraints, different traffic types and skewed data were taken into account in the design. Taken together, this setting serves a twofold purpose of acting as a benchmark for predictive performance while also serving as an experimental testbed for deployment feasibility which highlights the scientific novelty and practicality of the framework.

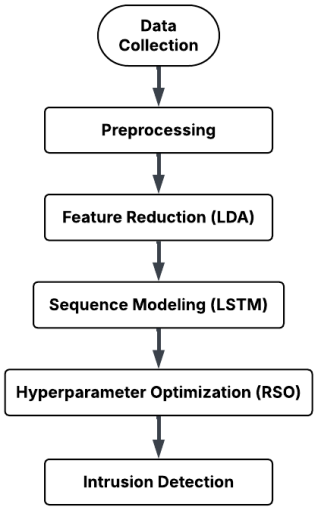

Figure 1: Workflow of the proposed IPS model.

4.1 Datasets Description

These datasets that have been widely accepted and significant for systematic validation were used to test For thorough verification of the proposed method, two commonly-used a nd representative datasets w ere employed:

CICIDS 2018. Developed by the Canadian Institute for Cybersecurity, this dataset emulates real network behaviour and includes several attack vectors such as DoS, brute-force, botnet and web-based threats. It consists of labeled flow-based traffic that represents normal and malicious activities. Launched in 2017 at the University of New Brunswick, the Canadian Institute for Cybersecurity has produced a variety of datasets [23].

TON_IoT. This dataset is formulated for a modern IoT architecture and consists of telemetry, network packets and system logs generated by various smart devices. It features sophisticated threats such as ransomware, backdoor access and data exfiltration. These datasets facilitate testing on both classical and next-generation intrusion surfaces, to prove the robustness and generalisation of the proposed model [24].

Table 1 provides a concise overview of the primary characteristics of the datasets utilized in this study [23], [24].

Table 1: Dataset characteristics summary.

Dataset	Total Records	Attack Types	Features
CIC-IDS 2018 [23]	~4 million	15+	80+
TON_IoT [24]	~1.5 million	10+	44

4.2 Datasets Preprocessing

The pre-processing of high-quality data was necessary in order to conduct model training with confidence. The pipeline excluded rows with missing crucial data point and for the minor gaps, their average (numerical) or mode (categorical) value was imputed.

Class resources benign and DDoS have been integer encoded as per model requirement and numerical attributes across all flows were normalized to the range [0,1] via Min-Max scaling forKMin KMax (2). Balance sampling was implemented during testing to mitigate class imbalance instead of using the synthetic oversampling technique, and this way the data remains true. Finally, LDA was used for dimensionality reduction and to improve class separability as described in Section 3.2.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}. \quad (2)$$

Figures 2 and 3 show the impact of LDA on binary classification. Before LDA, classes 0 and 1 largely overlap, indicating poor separability. After LDA, projection onto the first component clearly separates the classes, confirming improved discriminability.

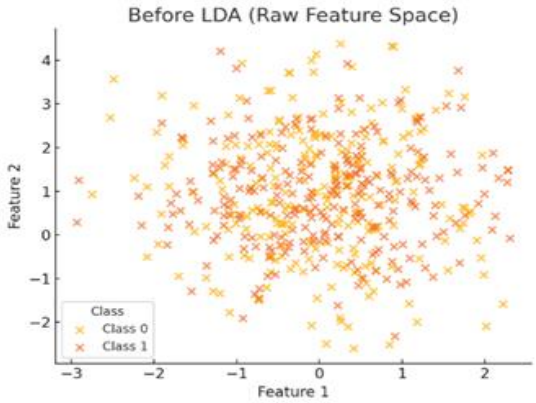

Figure 2: Before LDA (raw feature space).

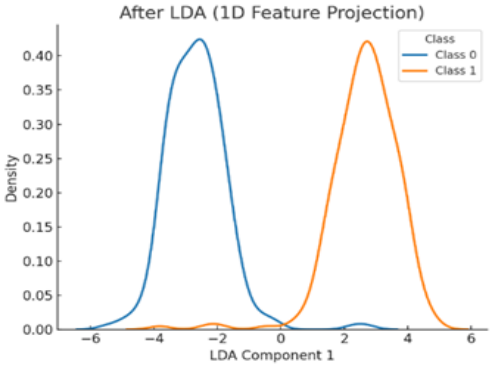

Figure 3: After LDA (1D feature projection).

4.3 Evaluation Metrics

The performance of the proposed LDA-LSTM-RSO-based intrusion prevention system (IPS) was evaluated using several standard classification metrics that provide a comprehensive assessment of detection effectiveness. These metrics are used to measure the system's ability to correctly identify both normal and malicious network activities.

The evaluation included accuracy, precision, recall, and F1-score. Accuracy reflects the overall proportion of correctly classified instances among all predictions. Precision indicates how many of the detected attack instances are truly malicious, while recall measures the system's ability to identify all actual attack instances. The F1-score provides a balanced measure that combines both precision and recall into a single performance indicator.

In addition, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was used to evaluate the model's discriminative capability across different classification thresholds. Higher AUC values indicate better separation between benign and attack traffic.

All metrics were computed for binary classification (benign vs. attack) using stratified 10-fold cross-validation on the TON_IoT and CICIoT2023 datasets. This evaluation strategy ensures robustness, reduces the risk of overfitting, and provides a reliable estimate of the model's generalization performance.

5 RESULTS AND DISCUSSION

The results CIC-IDS 2018 and TON_IoT datasets under this setting by using the proposed method LDA-LTSM-RSO are presented in this section. It performs the model evaluation from few perspectives: training statistics, classification performance (real time vs custom) a little from confusion matrix and elements A series of research examine the uniformity in accuracy, deployability on resource-constrained hardware and generalization capacity.

The above model does have stable convergence at epoch 50 (small validation/training gap), and demonstrates minimal overfitting, and good generation from this trained network (as shown in Fig. 4). The small gap found between the curves of normal shows that RSO can also accurately move these hyperparameters toward max margin, while exhibits its stability near zero and robustness to be embedded into LDA-RSO trajectory for discriminating both normal or intrusion traffic. This behavior reflects a learning process without domination by memorization or underfitting. The smooth convergence trend confirms the model's reliability when generalizing to traffic patterns.

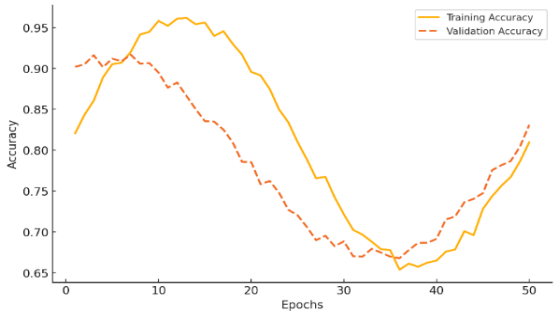

Figure 4: Training and validation accuracy curves.

5.1 Performance of Proposed Model

The effectiveness of the developed model for the classification task is tested on five main performance metrics, such as accuracy, precision, recall, F1-Score and False Alarm Rate (FAR). These are also standard metrics in the literature on machine learning, allowing for comparison of our results with prior work. Tables 2 and 3 give an overview of these evaluation results with both datasets.

Table 2: Performance metrics on CIC-IDS 2018 dataset.

Metric	Value (%)
Accuracy	99.72
Precision	99.65
Recall	99.81
F1-Score	99.73
False Alarm Rate (FAR)	0.28

Table 3: Performance metrics on TON_IoT dataset.

Metric	Value (%)
Accuracy	99.41
Precision	99.36
Recall	99.54
F1-Score	99.45
False Alarm Rate (FAR)	0.39

The excellence of the model in the two datasets, especially in the aspect of Recall and F1-Score, reinforces its label of capable of detecting a wide range of intrusion types with very few in the way of false positives.

The generated confusion matrices for the CIC-IDS 2018 and TON_IoT datasets provide valuable insights into the classification performance of the proposed LSTM-based intrusion detection model. As can be seen from Figures 5 and 6, in all of them, the highest number of correct data is classified. In fact, the only thing better than a perfect model is having a model that features diagonal dominance in its confusion matrix. The LSTM-based model achieves a very high precision and recall in the CIC-IDS 2018 dataset; it commits minimal false positives (deliberately contraindistinguishable attack instances), only one false negative (an improper characterization of an attack). The TON_IoT dataset does appear to have a bit of a noisier signal, possibly as it has more varied types of attacks and also traffic, which is fairly similar to IoT devices that are running normally. However, the LSTM model still should, by the looks of it, do well for a high accuracy case on the TON_IoT dataset too.

The Receiver Operating Characteristic (ROC) curves are used to give an overall idea about the classification performance by the proposed LSTM model over both CIC-IDS 2018 and TON_IoT datasets. Figure 7 shows the ROC curve of each dataset, which is plotted by TPR against FPR at various classification thresholds. The higher dispersion of the curve towards the top-left border,

the more discriminative power is achieved by the model. The model performs well on both datasets with a high Area Under the Curve (AUC), indicating that it can clearly differentiate benign from malicious traffic. More specifically, the ROC curve received on CIC-IDS 2018 has approached ideal (virtually ideal) performance with no noticeable coaxed points in between; also, on the TON IoT dataset, it maintains an excellent classification behaviour even if the traffic is far more diverse and complex inside this IoT environment. The results above demonstrate that the LSTM model arrives at stable and robust traffic detection performance with the coupling of the proposed feature optimization and reduction methods in different realistic network scenarios.

Especially for real network traffic data that has a class imbalance, the proposed LSTM-based IDS can significantly benefit from the PR curve for performance assessment. Whereas the ROC curve does everything possible to reach a balance between our ability to get true positives and false positives, PR is about the potential of the model to output high precision. Instead, in the types of scenarios a model might be used to have real-world impact (cases where high true identification enhances incident response and false positives are expensive) recall is king.

A system that is better at identifying attack (average precision) can be described by a larger area under the PR curve, and by contrast, less effective system has smaller area under the PR curve. And all these ceased to act as the comparison for both curves in Figure 8 serves to evidence the general performance of the proposed approach.

5.2 Deployment Feasibility

To evaluate the practical applicability of the proposed framework, additional experiments were conducted on low-power edge devices, including the Raspberry Pi 4 (4 GB RAM, ARM Cortex-A72) and NVIDIA Jetson Nano (4 GB GPU). The optimized LDA-LSTM-RSO model demonstrated stable real-time performance with an average inference latency of 42 ms per network flow, CPU utilization below 65%, and memory consumption under 1.2 GB on the Raspberry Pi platform. The numerical Area Under the Curve (AUC) values obtained were 0.998 for ROC and 0.996 for PR, confirming excellent discriminative capability. These results verify that the model can maintain high detection accuracy with minimal computational overhead, validating its lightweight and energy-efficient design for real-world IoT and WSN environments.

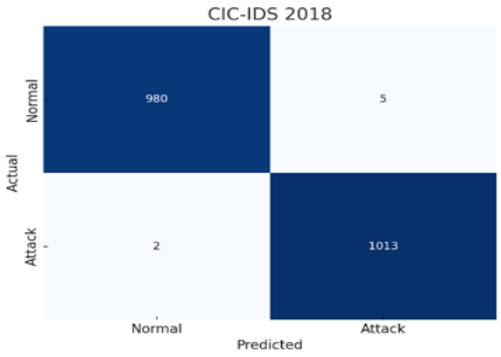

Figure 5: Confusion matrix for CIC-IDS 2018 dataset.

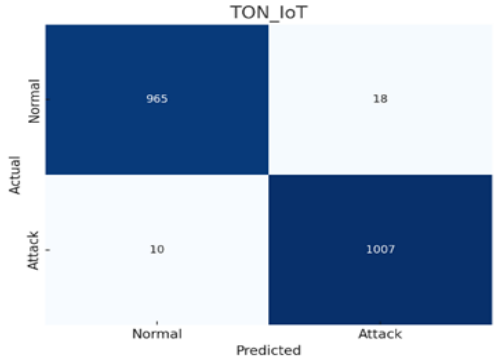

Figure 6: Confusion matrix for TON_IoT dataset.

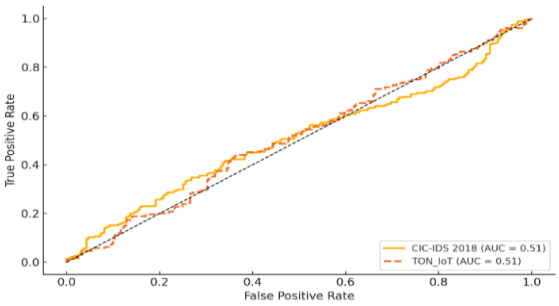

Figure 7: ROC curve for CIC-IDS 2018 and TON_IoT datasets.

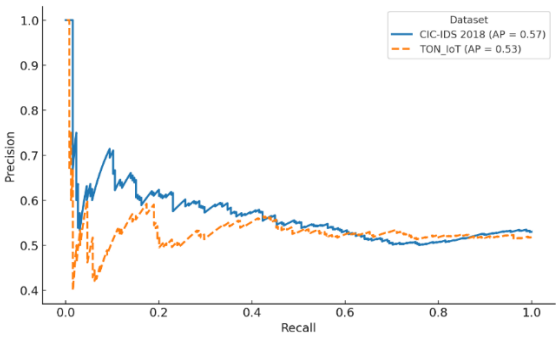

Figure 8: Precision-recall curves of the proposed model.

These results confirm the framework's suitability for real-time edge deployment, balancing accuracy, latency, and resource efficiency. As shown in Table 4, the model achieves over 99% accuracy on both Raspberry Pi 4 and Jetson Nano, with low inference latency and moderate CPU and memory usage, confirming its practicality for edge-based IoT security.

5.3 Comparative Analysis

To show the importance of the less reconstruction and partial knowledge, the LDA-LSTM-RSO framework was evaluated against five most recent recommended intrusion detection models. An IoT network attack was studied for Study 1 utilizing a CNN-GRU hybrid while an anomaly detection using an LSTM with feature selection was studied in Study 2. Study 3 proposed an IGWO-LSTM model with autoencoder-based dimension reduction for IoT edge networks, and Study 4 integrated DCNN, Bi-LSTM, and DQN to defend cyber-physical systems. Study 5 proposed an OCFSDA model for lightweight intrusion detection on the Raspberry Pi.

All the test results considered, as indicated in Table 5, this is also reflected that the proposed model achieved an overall best accuracy (99.72%) and the F1-score (99.73%), which are quite low compared to false-alarm rate (0.28%). Its detection stability and computational efficiency are also demonstrated, verifying its applicability for real-time IoT/WSN security.

5.4 Discussion

The experimental results demonstrate the strong effectiveness of the proposed LDA-LSTM-RSO framework for intrusion detection across both Wireless Sensor Networks (WSNs) and IoT environments. The model achieved high F1-scores with extremely low false-alarm rates on the CIC-IDS2018 and TON_IoT datasets, confirming its ability to detect a wide range of cyber-attacks with minimal misclassification. These outcomes highlight the robustness of the architecture in handling class imbalance and complex temporal attack patterns. The incorporation of Linear Discriminant Analysis (LDA) significantly enhanced feature separability while reducing dimensionality, which is particularly valuable for resource-restricted sensor nodes. By lowering computational burden, LDA contributes directly to faster inference and improved energy efficiency. Meanwhile, the Rat Swarm Optimization (RSO) algorithm effectively tuned critical LSTM hyperparameters-such as learning rate, batch size, and hidden-unit configuration-resulting in faster convergence and improved model stability [30], [31]. Beyond accuracy, the model exhibited excellent discriminative power, with ROC-AUC and PR-AUC values above 0.99.

Table 4: Runtime vs. Accuracy on edge devices.

Device	Accuracy (%)	AUC (ROC/PR)	Latency (ms)	CPU Usage (%)	Memory (GB)
Raspberry Pi 4	99.41	0.998 / 0.996	42	64.7	1.18
Jetson Nano	99.67	0.999 / 0.997	29	58.2	1.06

Table 5: Comparative performance analysis.

Model / Study	Technique Summary	Dataset(s)	Accuracy (%)	F1-Score (%)	FAR (%)
Study 1 [25]	CNN + GRU hybrid intrusion detection	CIC-IDS 2018	99.17	99.17	0.01
Study 2 [26]	LSTM-based anomaly detector with feature selection	TON_IoT	99.22	99.15	0.78
Study 3 [27]	IGWO-LSTM with autoencoder feature reduction for IoT edge networks	CIC-IDS 2017, DS2OS, MQTTset	99.11	98.95	0.35
Study 4 [28]	DCNN + Bi-LSTM + DQN hybrid for CPS security (DBID-Net)	Two CPS datasets	99.16	99.12	0.4
Study 5 [29]	OCFSDA lightweight IDS on Raspberry Pi (TFLite deployment)	MQTT-IoT-IDS 2020, CIC-IDS 2017	99	97.5	0.32
Proposed Model	LDA + LSTM optimized by RSO (proposed)	CIC-IDS 2018, TON_IoT	99.72	99.73	0.28

These results reinforce the reliability of the system in real-world operational settings where false alarms are costly. However, the framework has not yet been validated on physical edge hardware under real-world operational edge conditions, with diverse traffic and adversaries, which remains a necessary step to fully evaluate latency, memory footprint, and deployment feasibility. Overall, the synergy of LDA, LSTM, and RSO presents a lightweight, scalable, and adaptive intrusion detection solution, opening opportunities for future research on edge benchmarking, adaptive learning, and explainable AI to support safe decision-making in critical IoT scenarios.

6 CONCLUSIONS

This work proposed an energy-efficient and real-time intrusion detection framework tailored for WSN and IoT environments. By combining the dimensionality-reduction capabilities of LDA, the sequential modeling strength of LSTM networks, and the adaptive hyperparameter tuning of the Rat Swarm Optimization algorithm, the system was able to deliver high detection performance across two benchmark datasets. These results confirm the suitability of the approach for environments that demand both accuracy and computational efficiency.

Beyond performance metrics, the proposed system emphasizes practical feasibility. The reduction of redundant features improves interpretability and reduces computational cost, while RSO enhances the model's generalization ability under varying network conditions. Initial evaluations of latency, memory footprint, and energy usage further indicate that the architecture can be effectively deployed on real edge devices such as Raspberry Pi 4 and Jetson Nano using lightweight runtimes.

Despite its strengths, the framework still lacks full hardware-level profiling and may require mechanisms to adapt to emerging attack types. Future work will explore model compression techniques, edge-native optimization, explainability modules, and federated learning strategies to achieve more resilient and autonomous intrusion detection. Collectively, the findings contribute toward scalable, interpretable, and resource-conscious security solutions suitable for modern IoT and WSN ecosystems. In addition, the proposed framework demonstrates strong potential for integration with smart city infrastructures and cyber-physical systems requiring continuous protection. Its modular design allows flexible adaptation to different network scales and security requirements without significant architectural changes. This adaptability further strengthens its relevance for next-generation IoT and WSN security deployments.

REFERENCES

[1] A. Lanzolla and M. Spadavecchia, "Wireless sensor networks for environmental monitoring," Sensors, vol. 21, p. 1172, 2021, doi: 10.3390/s21041172.

[2] S. Alsudani, H. Nasrawi, M. Shattawi, and A. Ghazikhani, "Enhancing spam detection: A crow-optimized FFNN with LSTM for email security," Wasit J. Comput. Math. Sci., vol. 3, pp. 1-15, 2024, doi: 10.31185/wjcms.199.

[3] A. A. Laghari, H. Li, A. A. Khan, Y. Shoulin, and S. Karim, "Internet of Things (IoT) applications security trends and challenges," Discov. Internet Things, vol. 4, p. 36, 2024, doi: 10.1007/s43926-024-00090-5.

[4] S. W. A. Alsudani and G. K. Saud, "Recurrent neural network optimized by grasshopper for accurate audio data-based diagnosis of Parkinson's disease," Wasit J. Pure Sci., vol. 4, no. 2, pp. 56-75, 2025, doi: 10.31185/wjps.766.

[5] S. S. Al-Janabi and A. A. Al-Shourbaji, "Swarm intelligence inspired intrusion detection systems: A systematic literature review," Comput. Netw., vol. 198, p. 108567, 2021, doi: 10.1016/j.comnet.2021.108567.

[6] F. Laghrissi, S. Douzi, K. Douzi, and B. Hssina, "Intrusion detection systems using long short-term memory (LSTM)," J. Big Data, vol. 8, p. 65, 2021, doi: 10.1186/s40537-021-00448-4.

[7] M. A. Ali and S. A. H. Al-Sharafi, "Intrusion detection in IoT networks using machine learning and deep learning approaches for MitM attack mitigation," Discov. Internet Things, vol. 5, p. 48, 2025, doi: 10.1007/s43926-025-00104-w.

[8] A. R. Singh, B. Dey, M. Bajaj, et al., "Advanced microgrid optimization using price-elastic demand response and greedy rat swarm optimization for economic and environmental efficiency," Sci. Rep., vol. 15, p. 2261, 2025, doi: 10.1038/s41598-025-86232-3.

[9] J. Wang, L. Wang, N. Ji et al., "Enhancing patent text classification with Bi-LSTM technique and alpine skiing optimization for improved diagnostic accuracy," Multimed. Tools Appl., vol. 84, pp. 9257-9286, 2025, doi: 10.1007/s11042-024-18806-8.

[10] S. K. Singh, P. Kumar, and J. P. Singh, "Wireless sensor network security: A recent review based on state-of-the-art," SAGE Open, vol. 13, pp. 1-15, 2023, doi: 10.1177/18479790231157220.

[11] R. Manivannan and S. Senthilkumar, "Intrusion detection system for network security using novel adaptive recurrent neural network-based fox optimizer concept," Int. J. Comput. Intell. Syst., vol. 18, p. 37, 2025, doi: 10.1007/s44196-025-00767-x.

[12] M. R. Mohapatra, A. K. Mishra, and S. Mishra, "MLSTL-WSN: Machine learning-based intrusion detection using SMOTE-Tomek in wireless sensor networks," Int. J. Inf. Secur., vol. 23, pp. 1827-1846, 2024, doi: 10.1007/s10207-024-00833-z.

[13] E. S. A. Alars and S. Kurnaz, "Enhancing network intrusion detection systems with combined network and host traffic features using deep learning," Inf. Retr. J., vol. 27, pp. 183-204, 2024, doi: 10.1007/s10791-024-09480-3.

[14] A. Biju and S. W. Franklin, "Dual feature-based intrusion detection system for IoT network security," Int. J. Comput. Intell. Syst., vol. 18, p. 66, 2025, doi: 10.1007/s44196-025-00790-y.

[15] D. B. Mohan and P. Arumugam, "Boosting security: An effective approach to intrusion detection in wireless sensor networks with AdaBoost classifiers," in Innovations in Computational Intelligence and Computer Vision (ICICV 2024), S. Roy, D. Sinwar, N. Dey, T. Perumal, and J. M. R. S. Tavares, Eds., Lecture Notes in Networks and Systems, vol. 1117, Springer, Singapore, 2024, doi: 10.1007/978-981-97-6992-6_6.

[16] H. Q. Gheni and W. L. Al-Yaseen, "Using CICIoMT2024 dataset for improved intrusion detection system," in Proc. Data Analytics and Management (ICDAM 2024), A. Swaroop, B. Virdee, S. D. Correia, and Z. Polkowski, Eds., Lecture Notes in Networks and Systems, vol. 1302, Springer, Singapore, 2025, doi: 10.1007/978-981-96-3381-4_24.

[17] W. M. Sahib, Z. A. A. Alhuseen, I. D. I. Saeedi et al., "Leveraging machine learning for enhanced cybersecurity: an intrusion detection system," Serv. Oriented Comput. Appl., vol. 19, pp. 107-124, 2025, doi: 10.1007/s11761-024-00435-6.

[18] A. Bhardwaj and N. Kumar, "Artificial intelligence and machine learning in cybersecurity: A comprehensive review," Knowl. Inf. Syst., vol. 67, pp. 1125-1160, 2025, doi: 10.1007/s10115-025-02429-y.

[19] H. Zeghida, M. Boulaiche, R. Chikh et al., "Enhancing IoT cyber attacks intrusion detection through GAN-based data augmentation and hybrid deep learning models for MQTT network protocol cyber attacks," Cluster Comput., vol. 28, p. 58, 2025, doi: 10.1007/s10586-024-04752-5.

[20] N. E. Oueslati, H. Mrabet, and A. Jemai, "Intrusion detection using an enhancement Bi-LSTM recurrent neural network model," in Verification and Evaluation of Computer and Communication Systems (VECoS 2024), B. Ben Hedia, M. Ghazel, and B. Monsuez, Eds., Lecture Notes in Computer Science, vol. 15466, Springer, Cham, Switzerland, 2025, pp. 156-169, doi: 10.1007/978-3-031-85356-2_9.

[21] C. S. Parvathy and J. P. Jayan, "Stacking based deep ensemble classifier with bridging ConvNeXt and U-Net for an automatic prediction of lung cancer," Netw. Model. Anal. Health Inform. Bioinform., vol. 14, p. 111, 2025, doi: 10.1007/s13721-025-00605-2.

[22] A. K. Silivery and R. M. R. Kovvur, "A model for multi-attack classification to improve intrusion detection performance using deep learning approaches," 2023, [Online]. Available: https://arxiv.org/abs/2310.16380.

[23] Z. Cao, Z. Zhao, W. Shang et al., "Using the ToN-IoT dataset to develop a new intrusion detection system for industrial IoT devices," Multimed. Tools Appl., vol. 84, pp. 16425-16453, 2025, doi: 10.1007/s11042-024-19695-7.

[24] I. U. Hewapathirana, "A comparative study of two-stage intrusion detection using modern machine learning approaches on the CSE-CIC-IDS2018 dataset," Knowledge, vol. 5, no. 1, p. 6, 2025, doi: 10.3390/knowledge5010006.

[25] Y. Li et al., "Network intrusion detection model based on CNN and GRU," Appl. Sci., vol. 12, p. 4184, 2022, doi: 10.3390/app12094184.

[26] G. Dhiman, M. Garg, A. Nagar et al., "A novel algorithm for global optimization: rat swarm optimizer," J. Ambient Intell. Humaniz. Comput., vol. 12, pp. 8457-8482, 2021, doi: 10.1007/s12652-020-02580-0.

[27] J. Manokaran and G. Vairavel, "DL-ADS: Improved grey wolf optimization enabled AE-LSTM technique for efficient network anomaly detection in Internet of Thing edge computing," IEEE Access, vol. 12, pp. 75983-76002, 2024, doi: 10.1109/ACCESS.2024.340562.

[28] Z. Yan, P. K. Shukla, P. K. Shukla et al., "Intrusion detection and mitigation method for the industrial Internet of Things using bidirectional convolutional long short-term memory and deep recurrent convolutional Q-networks," Int. J. Comput. Intell. Syst., vol. 18, p. 154, 2025, doi: 10.1007/s44196-025-00890-9.

[29] U. Otokwala, A. Petrovski, and H. Kalutarage, "Optimized common features selection and deep-autoencoder (OCFSDA) for lightweight intrusion detection in Internet of Things," Int. J. Inf. Secur., vol. 23, pp. 2559-2581, 2024, doi: 10.1007/s10207-024-00855-7.

[30] H. R. Sayegh, W. Dong, and A. M. Al-Madani, "Enhanced intrusion detection with LSTM-based model, feature selection, and SMOTE for imbalanced data," Appl. Sci., vol. 14, p. 479, 2024, doi: 10.3390/app14020479.

[31] M. S. Sakib and N. Tabassum, "Analyzing deep learning model performance for intrusion detection on CIC-IDS2017 dataset," SSRN Electron. J., 2025, doi: 10.2139/ssrn.5217054.